

Core Information Management Standards for the EMAP Western Pilot Study

Introduction

The core data management and GIS standards for the Environmental Monitoring and Assessment Program (EMAP) Western Geographic Study are set forth below. This is not a comprehensive list of all federal or EPA standards, nor is it a list of data standards. It is a list of standards pertaining to information management that all participants in the EMAP Western Pilot Study agree to follow. The Landscapes, Streams, and Coastal groups will develop more detailed standards that are specific to their groups. Further details on EMAP standards are given in the EMAP Information Management Plan (Hale et al. 1999).

The goal of these core standards is to maximize the ability to exchange data within the Study and with other studies conducted under the monitoring framework of the Committee on Natural Resources and Environment (CENR 1997). The main standards are those of the Federal Geographic Data Committee (FGDC 1999), the National Spatial Data Infrastructure (NSDI 1999), and the National Biological Information Infrastructure (NBII 1999).

Metadata

- Federal Geographic Data Committee, Content standard for digital geospatial metadata, version 2.0. FGDC-STD-001-1998 (FGDC 1998), including the Biological Data Profile and the Biological Names and Taxonomy Data Standards developed by the National Biological Information Infrastructure (NBII 1999).
- For tabular data, metadata that meet the FGDC content standard are contained by a combination of the EMAP Data Directory and the EMAP Data Catalog. For Arc/Info coverages, the metadata are in the .DOC file embedded in the coverage. This file stays with the coverage. When the coverage is moved to the EMAP public web site, it will be duplicated to an ASCII text file.

EMAP Data Directory

- EMAP Data: EMAP Data Directory Oracle database. The guidelines are given in Frithsen and Strebel (1995), Frithsen (1996a,b), and USEPA (1996b).
- Other data: Environmental Information Management System (EIMS 1999). EMAP Directory entries are periodically uploaded to the EIMS. The EIMS will become EPA's node for the National Spatial Data Infrastructure and will make directory information available to other

federal agencies through the Z39.50 protocol in accordance with the US Global Change Research Program (USGCRP 1998).

EMAP Data Catalog

- EMAP Data Catalog standards are given in Strebel and Frithsen (1995b), Frithsen (1996a), and USEPA (1996c).

Data Formats

- **Attribute data**
 - ASCII files: comma-separated values, or space-delimited, or fixed column.
 - SAS export files.
 - Oracle
- **GIS data**
 - ARC/INFO export files; compressed .tar file of ARC/INFO workspace
 - Spatial Data Transfer Standard (SDTS; FGDC 1999) format available on request

Parameter Formats

EMAP has standard formats only for place and time.

- **Sampling Site** (EPA Locational Data Policy; EPA 1991)
 - Latitude and Longitude in decimal degrees (+/- 7.4)
 - Negative longitude values (west of the prime meridian).
 - NAD83
- **Date:** YYYYMMDD (year,month,day).
- **Hour:** HHMMSS (hour,minute,second)
 - Greenwich mean time
 - Local time

Data loaded to STORET will take on the STORET formats upon loading.

Standard Coding Systems

- **Chemical Compounds:** Chemical Abstracts Service (CAS 1999)
- **Species Codes:** Integrated Taxonomic Information System (ITIS 1999).
- **Land cover/land use codes:** Mutli-Resolution Land Characteristics (MRLC 1999)

EMAP Web Site and Data Distribution

- EMAP-funded data, directory entry and catalog files must be made available to the EMAP public web site (EMAP 1999).
- Guidelines for making data available on this site are given in Strebel and Frithsen (1995a) and (USEPA 1997).
- Data and metadata files are posted to the internal EMAP web site [<http://naraxp.nar.epa.gov/emap>] for review by the contributor before moving to the EMAP public web site (EMAP 1999).
- Data on the internal web site may have not gone through full Quality Control/Quality Assurance (QA/QC), but data to be placed on the public web site must have undergone QA/QC according to an approved Quality Assurance Project Plan.
- No data sent to EMAP public web site without approval from data source.
- No data sets are distributed publicly without the accompanying metadata.
- Web site design must follow EPA standards [<http://www.epa.gov/epahome/webguide/guide.htm>].
- Very large files may be distributed by CD-ROM or DVD (free or for no more than the cost of reproduction)

EMAP Bibliography

- EMAP Bibliography, an Oracle database (EMAP 1999).
- 'Guide to Submitting Information to the EMAP Bibliographic Database' (EMAP 1999).
- The citation format for the EMAP Bibliography is the Council of Biology Editors Manual (CBE 1994).

Data Stewardship and Responsibility

- Data collectors are responsible for the preparation of data, directory, and catalog files.
- Data stewardship, maintenance and implementation of data Quality Control/Quality Assurance procedures lies primarily with the data collectors.

Long-term Archival

- EMAP web data: EMAP Archival Plan (Hale et al. 1999; USEPA 1996a).
- Low-level data not transferred to EMAP web site: relevant EPA and Division archival policies.
- Water quality data will be archived to the extent feasible in STORET (STORET 1999).

References

CAS. 1999. Chemical Abstracts Service web site. [<http://www.cas.org>]

CBE. 1994. Scientific Style and Format. The CBE Manual for Authors, Editors, and Publishers. 6th ed. The Council of Biology Editors, Chicago, IL.

CENR. 1997. Integrating the Nation's Environmental Monitoring and Research Networks and Programs: A Proposed Framework. Committee on Environment and Natural Resources, National Science and Technology Council, Washington, DC, USA.

EIMS. 1999. Environmental Information Management System (EIMS) web site. [<http://www.epa.gov/eims>].

EMAP. 1999. Environmental Monitoring and Assessment Program (EMAP) web site. [<http://www.epa.gov/emap>].

FGDC 1998 Content standard for digital geospatial metadata, version 2.0. FGDC-STD-001-1998. Federal Geographic Data Committee, Washington, DC [<http://www.fgdc.gov>].

Frithsen, J. B. 1996a. Suggested modifications to the EMAP data set directory and catalog for implementation in US EPA Region 10. Draft, June 10, 1996. Report prepared for the U.S. Environmental Protection Agency, National Center for Environmental Assessment, Washington, DC., by Versar, Inc., Columbia, MD.

Frithsen, J. B. 1996b. Directory Keywords: Restricted vs. unrestricted vocabulary. Draft, May 21, 1996. Report prepared for the U.S. Environmental Protection Agency, National Center for Environmental Assessment, Washington, DC., by Versar, Inc., Columbia, MD.

Frithsen, J. B., and D. E. Strebel. 1995. Summary documentation for EMAP data: Guidelines for the information management directory. 30 April 1995. Report prepared for U.S. Environmental Protection Agency, Environmental Monitoring and Assessment Program (EMAP), Washington, DC. Prepared by Versar, Inc., Columbia, MD.

Hale, S., J. Rosen, D. Scott, J. Paul, and M. Hughes. 1999. EMAP Information Management Plan: 1998–2001. U.S. Environmental Protection Agency, Office of Research and Development, Research Triangle Park, NC.

ITIS. 1999. Integrated Taxonomic Information System web site [<http://www.itis.usda.gov/itis>].

MRLC. 1999. Mutli-Resolution Land Characteristics web site [<http://www.epa.gov/mrlc/>].

NBII. 1999. The NBII Biological Metadata Standard. National Biological Information Infrastructure web site. [<http://www.nbii.gov>].

NSDI. 1999. National Spatial Data Infrastructure web site [<http://www.fgdc.gov/nsdi/nsdi.html>].

STORET. 1999. The STORET web site. [<http://www.epa.gov/OWOW/STORET>].

Strebel, D. E., and J. B. Frithsen, 1995a. Guidelines for distributing EMAP data and information via the Internet. April 30, 1995. Prepared for U.S. Environmental Protection Agency, Environmental Monitoring and Assessment Program (EMAP), Washington, DC. Prepared by Versar, Inc., Columbia, MD.

Strebel, D. E., and J. B. Frithsen, 1995b. Scientific documentation for EMAP data: Guidelines for the information management catalog. Draft: April 30, 1995. Prepared for U.S. Environmental Protection Agency, Office of Modeling, Monitoring Systems and Quality Assurance, Washington, DC. Prepared by Versar, Inc., Columbia, MD.

USEPA. 1991. IRM Policy Manual. Chapter 13. Locational data.

USEPA. 1996a. Directive 2100-Information Resources Management, Chapter 10. Records management.

USEPA. 1996b. Addendum to: Guidelines for the information management directory. U.S. EPA, NHEERL, Atlantic Ecology Division, Narragansett, RI.

USEPA. 1996c. Addendum to: Guidelines for the information management catalog. U.S. EPA, NHEERL, Atlantic Ecology Division, Narragansett, RI.

USEPA. 1997. Update to: Guidelines for distributing EMAP data and information via the Internet. U.S. EPA , NHEERL, Atlantic Ecology Division, Narragansett, RI.

USGCRP. 1998. Data Management for Global Change Research. Policy Statements for the National Assessment Program. July 1998. U.S. Global Change Research Program. National Science Foundation, Washington, DC.